A Systematic Review of DTI studies in Bipolar Disorder

By Gareth Nortje

MB.ChB.(UCT), MSc.(Edinb), MRCPsych, FRCPsych(SA)

NRTGAR001

g.nortje@gmail.com

Dissertation submitted in partial fulfilment of

MMed(Psych) Part III

Supervisor: Neil Horn, MBChB, FCPsych(SA)

Department of Psychiatry and Mental Health

Faculty of Health Sciences

University of Cape Town

November 2012

Prepared in a format suitable for submission to *Bipolar Disorders*

Official journal of The International Society for Bipolar Disorders

http://www.blackwellpublishing.com/journal.asp?ref=1398-5647

Word count (excluding references, captions and supplementary material): 7275

Systematic Review of DTI studies in Bipolar Disorder

Gareth Nortje NRTGAR001

Department of Psychiatry and Mental Health

University of Cape Town

Abstract

Background: In the last decade, multiple diffusion tensor imaging (DTI) studies have revealed changes in the microstructure of white matter in bipolar disorder. The results are poorly replicated and inconsistent, however, with some authors suggesting a predominance of alterations in fronto-limbic white matter. Preliminary reading of the literature suggests that white matter changes as revealed by DTI may be more widespread throughout the brain. Two extant reviews have each been limited by including all affective disorders or by a methodology which ignores tracts and discards potentially meaningful data. This background in the review includes a detailed exposition of the main DTI techniques and shortcomings.

Aim: The review aims to determine whether certain white matter tracts are affected preferentially in the brain, as opposed to more diffuse white matter involvement. It also aims to determine if there is an anterior-posterior gradient of abnormalities.

Method: This review systematically collates data relating to tract involvement as demonstrated by DTI, as well as data regarding anterior-posterior distribution of abnormalities. Medline and EMBASE databases are searched systematically to select original papers comparing a bipolar group with healthy controls, using DTI, in adults, and reporting at least fractional anisotropy (FA). Subject, scan and analysis characteristics are extracted. Details of affected tracts are collated, as is the y-axis (anterior/posterior) of the most affected ("peak") voxels.

Results: Results are tabulated and represented graphically, displaying the distribution of tracts most commonly affected, as well as the anterior/posterior distribution in the brain.

Discussion: The discussion includes a qualitative summary of the findings. The difficulty of performing a meaningful statistical analysis are discussed. Limitations are also discussed.

BACKGROUND

Bipolar disorder (BD) is a common disabling condition characterised by episodes of mania, depression or mixed mood symptoms. The marked changes in mood, energy, motivation, judgement, attention and biological rhythms characteristic of mood episodes are accompanied by subtle but persistent inter-episode deficits in working memory, attention and impulsivity (1). The pathophysiology of BD remains poorly understood, though structural (2, 3) and functional (4) imaging have revealed regional brain differences compared to healthy controls. Abnormalities of frontal cortical control of limbic systems have been suggested (5), which may account for emotional dysregulation as well as some executive deficits. Subtle abnormalities of white matter may be responsible for this relative functional disconnection between cortical, limbic and other brain regions (6).

Diffusion-tensor imaging (DTI) is a modern extension of MRI technology particularly suited to examining the microstructure of white-matter, as it is sensitive to the direction and speed of water diffusion in tissues. By measuring the average diffusion of water in a voxel from at least six directions, diffusion-weighted imaging can work out the eponymous "tensor", a mathematical construct with six variables, for each voxel. From the diffusion-tensor, one can deduce the most common values reported in DTI studies. Mean diffusivity (mD) is a scalar value reflecting the mean amount of water diffusion in any direction within a voxel. Fractional anisotropy (FA) is a measure of the diffusion *anisotropy*, or the tendency of water to diffuse preferentially along one axis within a voxel. FA is a scalar value varying from zero - perfect *isotropy* as may be found in cerebrospinal fluid, to one - perfect *anisotropy*, which indicates diffusion in one direction only. An FA of one is thus a theoretical maximum and never realised in the brain. White matter (WM) typically has the highest anisotropy ranging from 0.3 to 0.8, as water diffuses more easily longitudinally along axons rather than across them (7). Grey matter usually has some degree of anisotropy, though less than WM. FA is the most commonly reported metric in DTI studies, as it is assumed to contain the most information about the microstructural integrity of WM. However, the exact interpretation of changes in FA remains unclear (8). Other metrics reported in DTI studies are axial and radial diffusivity, reflecting the degree of diffusion parallel to and perpendicular to the principle axis of diffusion respectively. These measures, especially when combined with FA and mean diffusivity, may aid in clarifying the mictrostructural tissue changes responsible for changes in FA (9, 10). Since diffusion in DTI is measured as the mean over a whole voxel, which may have a volume of 8-27 mm^3 and contain thousands of axons and extracellular material, reported FA depends on a number

of tissue microstructural features. Myelination, axon diameter, axon membrane integrity, oedema, fibre coherence or fibre crossing all contribute to anisotropy and changes in measured FA may be due to any combination of these factors (8, 11, 12).

DTI data may be analysed in different ways, each with their own strengths and weaknesses. The simplest method is a "region-of-interest" (ROI) approach, where DTI measurements such as FA are averaged over a predefined area in the brain. Thus, ROI methods are hypotheses driven, but ignore the rest of the brain, leading to potential confirmation bias. Results from ROI studies are also critically dependent on exact ROI definition and placement (13), making comparison between studies difficult. Analysing the whole-brain on a voxelwise basis is statistically more challenging, but these voxel-based analyses (VBA) have superseded ROI studies in BD. VBA studies report cluster of contiguous voxels which are statistically significantly different. Most VBA studies report the coordinates of a "peak-voxel" where the difference is greatest, and many studies label the tracts which most likely impinge on the affected clusters. In both ROI and VBA studies, it may be difficult to correctly identify or isolate changes to a specific tract, as clusters may be irregular in shape, be located in areas of many crossing tracts, or overlap more than one tract. Tract-based spatial statistics (TBSS) is a whole-brain analysis method which compares only the core of the white matter tracts, or skeleton, common to most of the subjects in the study (14). Results are thus reported as thin clusters of contiguous voxels located in the centre of white matter tracts. Labelling these tracts is often facilitated in this manner, but is still imperfect as a single white matter core may carry fibres from multiple different tracts. The final commonly used method of analysis is tractography, which defines a ROI along most of the volume of a tract, using the principle direction of diffusion in each voxel to track fibres probabilistically. (For more detailed discussions of the technology of DTI imaging and strengths and limitations of analysis methods, see the supplementary material.)

DTI has been used to examine possible WM changes which may contribute to the pathophysiology of BD. All four analysis methods have been applied, but results have generally been poorly replicated and inconsistent, with multiple authors attempting to review and synthesise the data.

Existing Reviews of the Literature

Vederine et al. (15) performed a meta-analysis of whole-brain voxel-based DTI studies using anatomical likelihood estimation (ALE) analysis to locate clusters of most consistently reported altered FA. Starting from the observation that peak voxels are merely estimates of the location of real peak differences between subjects, ALE takes the coordinates of a set of peak voxels and calculates in a statistically defensible manner the likelihood that they are randomly distributed in the brain (16). If the co-ordinates are not randomly distributed, ALE will generate cluster(s) indicating the likely underlying abnormality. For DTI studies of BD patients, Vederine et al. (15) analysed the peak voxel co-ordinates of clusters of decreased FA from ten whole-brain studies. Two significant clusters of decreased FA were identified, both on the right side. The first cluster was located close to the right parahippocampal gyrus posteriorly, with tractography from that cluster implicating various long association tracts. The second cluster was close to the right anterior and subgenual cingulate cortex, implicating the inferior fronto-occipital and uncinate fasciculi as well as forceps minor. Tractography from both clusters was performed on a single subject's DTI scan, and the authors recommend caution in interpreting the tractography results. In addition, ALE only takes account of the coordinates of the peak-voxel of each cluster, ignoring cluster size, shape, sign and absolute value.

Sexton et al. (17), in their systematic review of DTI studies in affective disorders (including major depression), found consistently reduced FA in frontal and temporal regions and tracts in patients relative to controls. They also performed a meta-analysis limited to nine papers with ROI placed "in or adjacent to the superior frontal gyrus". Results showed significant decreased FA in this region compared to healthy controls. Sexton et al. caution that the results and methodologies of the source studies were significantly heterogeneous, and demonstrated likely publication bias. Both Vederine et al. and Sexton et al. focussed on FA rather than other measures of diffusivity, as FA was routinely reported in contrast to other diffusion measures.

Heng et al. (18) did a qualitative systematic review of DTI studies in bipolar disorder including studies up to July 2009, noting that white matter changes are widespread but are most consistent, with decreased FA usually, in the frontal and prefrontal lobes. Brambilla et al. (19) in their discursive review, concluded that the evidence from diffusion imaging studies suggested impairments in fronto-limbic circuitry

(cingulum bundle and uncinate fasciculus), inter-hemispheric connectivity (corpus callosum) and fronto-parieto-temporal connections (superior longitudinal fasciculus).

Motivation for review

While limbic and frontal tracts in anterior WM are thought to be most affected in BD, in line with theories of emotion dysregulation (5, 20, 21), there has not yet been a systematic review which specifically examines whether these tracts are more consistently and more robustly affected than other tracts. Indeed, it is unclear whether certain classes of tracts – association, projection or commissural – are preferentially affected. ROI studies focussing on frontal WM (22-25) and corpus callosum (26-28) are over-represented, leading to potential confirmation bias as other WM areas are less studied. Knowledge of the regional distribution of WM changes, particularly if tract specific, may provide clues to their aetiology and pathogenesis. A further possibility is that there is anterior-posterior gradient of WM changes, affecting anterior brain regions more than posterior ones, as has been robustly demonstrated in normal aging (29). If this were so, then not only anterior tracts, but also anterior regions of long association and commissural tracts, would be preferentially affected.

In this systematic review we aim to identify which white-matter tracts, if any, are most consistently implicated by changes in anisotropy in bipolar disorder. We collate and summarise unbiased data without *a priori* assumptions from whole-brain VBA and TBSS studies. We supplement these findings with data from tractography studies, as although tractography does make *a priori* assumptions regarding areas studied, it has the best face validity regarding involvement of whole tracts. No previous reviews have systematically quantified tract involvement in BD. We also aim to demonstrate the presence or absence of an anterior-posterior gradient in anisotropy changes. Our main hypothesis is that certain tracts are implicated more frequently and consistently than others. Our secondary hypothesis is that anterior brain regions are affected more consistently than posterior. Both hypotheses are suggested by recent reviews of the DTI literature in bipolar disorder (15, 17). While we would have liked to investigate laterality, too few of the whole-brain studies (9 out of 15) limited their samples to right-handers to allow useful comment.

METHOD

Paper selection

The Pubmed/MEDLINE and EMBASE databases was systematically searched using the strategy (("bipolar" OR "mania" OR "depression") AND ("DTI" OR "diffusion tensor")) for papers published up to June 2012. In addition, reference lists of included studies were hand searched for suitable papers. Papers had to meet the following inclusion criteria: (1) be published in English in a peer-reviewed journal, (2) be about bipolar disorder, (3) contain original data from an original study (not a review), (4) compare a bipolar disorder group with a healthy control group (both bipolar I and II or unspecified were accepted), (5) subjects to be diagnosed using a validated diagnostic schedule according to DSM-IV or ICD-10 criteria, (6) use DTI as the imaging methodology, and report differences in fractional anisotropy, (7) subjects should be 18-65 years old.

As in previous reviews of DTI studies in bipolar disorder, we limited our analysis to the fractional anisotropy (FA) results, as other measures of diffusivity were reported too infrequently to allow useful comparisons. Abstracts were screened by two reviewers independently to assess for study inclusion, and differences were discussed and resolved by consensus.

Data capture

For the whole-brain VBA studies, each significant reported cluster was grouped according to the tract most likely involved, using the author's tract labels as supplied in each paper. Where more than one tract is named, the cluster was classified as involving *all* mentioned tracts (ie: one cluster can be included under more than one tract). Where only co-ordinates of peak-voxels are given without tract names, the co-ordinates will be entered into FSLView (www.fmrib.ox.ac.uk/fsl/fslview/) and labelled according to the supplied JHU White Matter Atlas. If the atlas cannot produce tract labels with at least 25% certainty for at least half the clusters in the study, the paper was excluded from the quantitative review. For the whole-brain TBSS studies, tract names as supplied were used. Where a single study used both TBSS and VBA analysis on the same subjects, the TBSS results took precedence as these are more likely to accurately identify tracts. Some papers report results both before and after correction for multiple comparisons, but only results which remain significant after correction are included.

For each affected tract in turn, we report how many studies report abnormalities in the tract, and the direction of the FA changes. In addition, following Schmahman and Pandya (30)(cited in (29)), we classify the tracts into their broad functional groupings (association, commissural, projection) and count how often these groups are reported to be affected. For the whole-brain studies we record the anterior-posterior co-ordinates, on the Talairach-Tournoux template, of the peak voxel of all significant clusters. For tractography studies we record which tracts were investigated and the significance and direction of differences in mean FA over the tract.

RESULTS

Paper selection

Systematic search of databases revealed abstracts of 208 English papers on 31 July 2012, which were reduced to 15 whole-brain studies using DTI to compare FA between bipolar patients and healthy controls in adults (Figure 1). Details of the 10 VBA studies (n=251) and the 5 TBSS studies (n=138) are shown in tables 1 and 2. Five papers used tractography in adults. Ha et al. (31) included only subgroup analyses of BDI and BDII patients compared to controls, of which we used the BDI results only as this subgroup is closer to the core bipolar disorder phenotype. Wessa et al. (32) do not include tract labels, and the coordinates of peak voxels are located peripherally in grey matter, making identification with FSLView impossible. These clusters are thus excluded from the tract analysis but included in the anterior-posterior analysis. Lu et al. (33) note that multiple association, commissural and projection tracts are affected in their study, but provide no labels or coordinates, so these results are excluded from both analyses, leaving 14 papers.

Subject and scan characteristics

Patient characteristics were heterogeneous along multiple dimensions, and no particular patient group could be considered wholly “typical”. Table 1 shows the clinical features of the subject groups and highlights where each group differs notably from the other studies. High heterogeneity of subject characteristics in terms of diagnosis, chronicity, clinical state, substance histories and genetic loading are apparent, each of which may influence DTI results. In all papers, patients were matched to healthy

controls by age and gender. Four papers stratified patient groups by lithium use (34), bipolar subtype (31, 35) or clinical state (36). For this review, the results from the non-stratified patient groups are used if available, as the combined groups are larger and more representative of the other papers' subjects. Versace et al. (37) and Zanetti et al. (36) used largely overlapping patient groups, but as each paper used a different analysis method both papers are included.

Scanning sequences in all studies were similar, though parameters such as voxel size, number of gradient directions and magnet strength differed between studies (Table 2). Notably, Bruno et al. (38) used much larger voxels which tend to report lower FA values in areas of crossing fibres and may be more sensitive to macroscopic differences in tract architecture rather than microscopic differences in axonal packing and myelination (39).

All of the fourteen whole-brain studies reported FA (as a requirement for inclusion in the review). In addition, five studies reported mean diffusivity, one apparent diffusion coefficient, four radial diffusivity, and three axial diffusivity. These supplementary measures of diffusivity, while informative about underlying pathology in individual studies, were inconsistently applied and did not allow for generalizable comment in this review. Five tractography studies met inclusion criteria (Table 3).

Cluster characteristics

The 14 whole-brain studies found a total of 51 significant clusters and 60 tract implications (Table 4, a single cluster could implicate more than one tract.). Despite using similar statistical analysis and thresholding, cluster characteristics varied greatly between studies. Each study contributed between one and eight clusters, though one study (40) using TBSS did not find a single significant cluster between bipolar subjects and healthy controls. Notably, this study was primarily comparing bipolar patients who did or did not have a history of suicide attempt, and the comparison with healthy controls was not the main aim. The largest clusters reported by each VBA study varied from 128 voxels to 13876 voxels (36, 41), a more than 10-fold difference, despite using the same voxelwise threshold for significance. Most studies (10/14) found only regions of decreased FA relative to controls, two studies found regions of both increased and decreased FA, and one study found only regions of increased FA. (For brevity and clarity, in this report all FA results are reported relative to the healthy control group. So "decreased FA" means "decreased FA *in the bipolar patients compared to healthy controls*".)

In the whole-brain analyses as a group there were more clusters of decreased FA than increased FA (34 vs 17, ns). However the TBSS studies were significantly more likely to find increased FA than the VBA studies ($p < 0.01$). TBSS studies showed a tendency to identifying more clusters than VBA studies.

Tracts implicated

All clusters from the nine VBA studies were supplied with tract labels and are included in the tract frequency analysis. Eleven VBA clusters implicated more than one tract. In the TBSS studies, two clusters from Chan et al. (42) and all seven clusters from Wessa et al. (32) did not have tract labels reported and FSLView was unable to supply labels from the peak-voxel co-ordinates. These nine unlabelled clusters are thus omitted from the tract frequency analysis. Notably, Wessa et al. was the only study to show increased FA relative to controls in all reported clusters.

In the VBA studies, tracts of all three main classes – commissural, association and projection – were implicated similarly, with most clusters demonstrating decreased FA (figure 2).

In the four TBSS studies, tracts of all three main classes are also implicated, though the pattern of tracts affected differs from the VBA studies (figure 3). The long association tracts (SLF, ILF, FOF) which are prominently represented in the VBA findings are not present in the TBSS, whereas the opposite is true for the posterior thalamic radiations. There may be artefactual reasons for these differences between VBA and TBSS findings which will be discussed later. TBSS identified a significantly greater proportion of areas of increased FA, and this finding would be even more significant had the seven unlabelled significant tracts of Wessa et al. (32) been included.

In the tractography studies, only certain tracts were investigated, so the absence of any entry on the chart merely signifies that there is no tractography data for that tract (figure 4). In the tracts which were investigated, findings of decreased FA were localised to uncinate fasciculus and anterior thalamic radiation, while other tracts found no significant differences.

Antero-posterior distribution of significant clusters

After conversion of all peak-voxel coordinates to Talairach Space, the y-values signifying anterior-posterior position were plotted as shown in Figure 5. The density of peak voxels cannot be inferred directly from the figure, as correction must be made for the volume of WM in which the voxels are found. If a tract were larger more anteriorly then even with uniform distribution of peak voxels one would expect to find more affected voxels anteriorly. To our knowledge there is no published data of mean WM or tract volumes over the antero-posterior extent of the brain. It is thus difficult to quantify the extent of possible A-P gradient in a statistically defensible manner. We can however make some observations.

After correcting for duplicate tracts implicated by the same cluster, there are 20 clusters anterior to the anterior commissure, 18 clusters between the anterior commissure and the posterior edge of the midline corpus callosum, and 13 clusters posterior to the corpus callosum. There does not appear to be any clear antero-posterior gradient of altered FA when peak voxels from the whole brain are considered together. The corpus callosum seems to have predominantly anterior involvement. Commissural fibres, projection fibres and association fibres are all implicated by alterations in FA. There does not appear to be a predilection for any particular class of tract.

DISCUSSION

Our analysis shows that all major classes of white matter tracts are implicated by alterations in FA in bipolar disorder. Furthermore, voxel-based and tract-based whole-brain analyses both implicate all classes of tract, though the pattern of specific tracts implicated differs by analysis method as shown in figures 2 and 3. Specifically, the long association tracts are implicated in VBA analyses but not by TBSS. The anterior and superior thalamic radiations are implicated by VBA, whereas TBSS rather finds significant alterations in the posterior thalamic radiations. The cingulum is implicated by half of TBSS studies but only 1/9 of VBA studies. Possible reasons for these discrepancies will be discussed below.

There are differences in how often individual tracts are implicated. For example, the ATR, SLF and other large tracts are each implicated by at least four papers, while the PCT is only implicated by one cluster in a single paper, and some important tracts, such as the arcuate fasciculus, are never implicated. This may be because the larger tracts are more severely affected by pathology in bipolar disorder, or it may be due to the confounding interactions of tract volumes, analysis method and labelling. First, voxel-based

analyses are usually thresholded to exclude smaller clusters, thus it is plausible that thinner smaller tracts would be less likely to achieve significance with an equivalent degree of FA alteration. This effect is exaggerated by partial-volume effects which affect thinner tracts especially (43). There is a close correspondence between how often tracts are implicated by VBA, and the volumes of selected tracts as measured using tractography (44). Second, a labelling bias might favour tracts which are larger and more anatomically consistent, as atlases tend to emphasize the constant or average anatomy of representative subjects at the expense of normal variability, especially of smaller tracts (45). During manual labelling, larger more conspicuous tracts in atlases may be favoured over smaller ones when cluster size and shape do not exactly match a known tract. It is not possible to correct for these confounders statistically, as reliable estimates of tract volumes for all major tracts are not yet available (44, 46), and labelling bias is subjective.

Thus although some of the larger tracts are implicated more frequently than smaller tracts, these findings are consistent with the hypothesis that white matter alterations are diffuse and widespread rather than tract-specific in bipolar disorder.

Even the most consistently implicated tracts are not robust and often demonstrate conflicting results. The most consistently implicated tract in VBA analyses, the anterior thalamic radiation (ATR) is implicated in five out of nine papers, but shows both increased and decreased FA. The ATR is one of the largest WM tracts (44), and with diffuse WM involvement one would expect it to be most affected. In the TBSS studies, the posterior thalamic radiation is most consistently implicated, but only in two out of four studies, each with opposite findings (37, 42). Lin et al. (21) hypothesised that the main tracts mediating frontal cortex connectivity (ATR, UF, SLF, cingulum, and inferior FOF) were preferentially affected in bipolar disorder, but demonstrated this only in the ATR and UF using tractographic analysis. In this review, the frontal cortex tracts hypothesised by Lin et al. are *all* implicated by at least two papers, though so too are the inferior longitudinal fasciculus and the superior and posterior thalamic radiations, which are not involved in frontal cortex connectivity (46). Similarly, Benedetti et al. (20) in their novel tractographic analysis of seven tracts between amygdala and other limbic areas, found decreased FA only in the fibres from SGC to amygdala (UF). (They did find other diffusivity abnormalities such as increased mean diffusivity in some other tracts, however.) No clear evidence for alterations of FA in frontal over non-frontal white matter tracts, nor indeed in limbic affective tracts over non-limbic tracts, has yet been demonstrated in bipolar disorder.

It is possible that alterations in diffusivity are widespread in the brain and not localised preferentially to specific tracts. The dependence of DTI results on issues of ROI placement suggests that underlying diffusivity abnormalities, if present, are likely to be subtle (Kanaan 2006). In the case of schizophrenia, a larger, well-powered VBA study (Kanaan 2009) revealed widespread diffusivity abnormalities which had been inconsistently and patchily revealed by smaller studies, leading the authors to suggest that WM changes may be widespread and subtle, accounting for the mixed findings of smaller studies. Diffuse rather than localised WM changes might be due to alterations in myelination genes, as has been posited in schizophrenia (Davis 2003) and bipolar disorder (Sprooten 2009).

Although frontal cortical tracts may not be specifically implicated when considered along their whole length, it is possible that anterior regions of white matter are affected more than posterior in bipolar disorder. Such an anterior-posterior gradient of changes in FA has been well demonstrated in normal aging (47, 48). However, our analysis of peak voxel co-ordinates revealed an even distribution from anterior to posterior. This even distribution held for projection and association fibres separately. However the corpus callosum, the main commissural bundle, seemed to be implicated more anteriorly than posteriorly, and demonstrated consistently decreased FA compared with controls (figure 5).

Overall this analysis of tract involvement suggests that WM abnormalities, as demonstrated by altered FA, are not reliably or consistently limited to particular tracts in bipolar disorder. The poorly replicated and varied findings of individual studies undermine confidence in their interpretation and their extrapolation to bipolar disorder patients in general. Whereas robust WM changes might be relatively impervious to study design, more subtle changes may interact with analysis technique and subject characteristics to influence the location, size, distribution and direction of abnormalities, as has been suggested in DTI studies of schizophrenia (Kanaan, 2006; Melonakos, 2011). In this review, the widespread FA changes coupled with inconsistencies between studies suggest that WM abnormalities, if present, are likely to be diffuse and subtle.

The difference in findings between VBA and TBSS analyses is notable. VBA analyses include all voxels of the white matter, including peripheral areas of architectural complexity and fibre crossings, while TBSS analyses consider only the central skeleton of tracts, where they are at their most compact and

coherent. Differences in findings between the two methods are thus not surprising, and may contain clues to the underlying pathology.

VBA studies tended to locate clusters of decreased FA in long association tracts which were not found with TBSS. However, in general TBSS analyses identified a significantly greater proportion of clusters with increased FA compared to VBA. This is an unexpected post-hoc finding, and needs to be replicated in different subject groups. For an example of this effect, Versace et al. (37) and Zanetti et al. (36) analysed data from largely overlapping subjects. Zanetti et al., using VBA, found clusters of decreased FA, whereas Versace et al., using TBSS, found clusters of mostly increased FA, in different areas. Note that Zanetti et al. did find clusters of increased FA when limiting their analysis to the acute depressed subgroup only.

Tract-specific differences in findings between different analysis methods may further illuminate the nature of the anatomical abnormalities. The cingulum, for example, was identified by two out of four TBSS papers with decreased FA, but was only implicated by one of the nine VBA papers, nor was any significant difference found in the single tractography study which included the cingulum. This suggests that there may be localised abnormalities in the core of the cingulum, which may be too thin be pick up on VBA studies, which typically have a cluster size-threshold of 50 voxels. Further, the non-significant tractography finding, which necessarily takes the mean over the majority of the tract, suggests that abnormalities may involve the tract unevenly. This is partly corroborated by the ROI study of Wang et al. (49) which found FA abnormalities in the anterior but not the posterior cingulum. Conflicts remain, however, as the TBSS studies reviewed here located abnormalities in the posterior cingulum rather than anterior.

TBSS and VBA analyses also implicate projection tracts in different areas of the brain. VBA studies found (conflicting) alterations in FA in the anterior projection fibres, while TBSS studies found alterations mostly in the posterior projection fibres. It is unclear why these methods locate FA alterations in different projection fibres, though it may be related to differences in tract thickness, coherence and inter-individual variability which affect registration differently between the two analysis methods (45).

The uncinate fasciculus, connecting orbitofrontal cortex to anterior temporal limbic structures, is the tract most studied by tractography in bipolar disorder, due to its prominent role in models of emotion

regulation (5). Of the four tractography studies involving the uncinate fasciculus, three found significantly decreased FA in bipolar compared to controls, though surprisingly this tract was only implicated by two VBA and one TBSS studies (31, 37, 41). In contrast to the tractography studies, the TBSS study found mostly, but not exclusively, increased FA. The uncinate fasciculus may not have been implicated in many whole-brain studies as it is relatively small in volume (44).

Our observation of significant differences between VBA and TBSS findings with respect to direction and location of FA changes, and the examples cited above, serve to emphasise that VBA and TBSS identify different aspects of white matter pathology. Opposite FA measurements between TBSS and VBA may not necessarily be conflicting, even in the same patient group, but should rather be viewed as complementary methods of analyses (36, 50).

Bipolar disorder is not a homogenous condition. Type I and Type II differ in genetic liability, gender distribution, clinical severity and presence of psychosis (51). Psychosis in bipolar disorder has genetic overlap with schizophrenia (52, 53), and may be associated with more severe neurocognitive deficits (54, 55). Familial liability to bipolar disorder has been associated with abnormalities in WM structure in the absence of clinical disorder (56, 57). Medication may have effects on DTI measures, though perhaps less than has been previously imagined (58). Length of illness, age, use of substances and current clinical state may all complicate DTI results (34, 47, 59).

Although most reviewed studies included only bipolar type I patients, four studies included mixed bipolar groups. Liu et al. (35) compared bipolar type I and type II in a subgroup analysis and found localised differences in FA in the ATR, anterior SLF and posterior cingulum. They suggest that there may be distinct neuropathological substrates for bipolar type I and type II, which would weaken or confound studies combining the two subtypes.

Two VBA studies limited their analyses to Type I bipolar patients with a 1st degree family history of bipolar disorder and a history of psychosis (41, 56). Both studies identified clusters of decreased FA in anterior association fibres and anterior corpus callosum, with Sussman et al. finding in addition decreased FA in regions of anterior and superior thalamic radiations. In addition, Chaddock et al. found that increasing genetic liability to bipolar disorder in unaffected relatives, as well as patients themselves,

correlated with decreased FA, suggesting that reduced FA might be a marker of genetic liability to bipolar disorder. Genetic liability thus becomes another source of heterogeneity in the reviewed studies.

Clinical state robustly affects functional connectivity between prefrontal and anterior limbic regions in bipolar patients (60), but it is unclear whether this is mediated by changes in white matter microstructure or synaptic transmission. Although changes in diffusivity as measured by DTI are thought to be primarily mediated by passive structural rather than active physiological characteristics, it is possible that localised inflammation with increased extra-cellular water content may alter diffusion characteristics relatively acutely (11). In addition, there is some evidence of minor hyperacute changes in mean diffusivity related to neuronal activation (61).

In their subgroup analysis comparing depressed and remitted patients, Zanetti et al. (36) found that abnormalities in FA were limited to the depressed subgroup. Zanetti et al. thus suggest that there are state-dependent microstructural white-matter changes in bipolar disorder. If this hypothesis is correct, it may be that some subtle state-dependent WM changes are diluted to non-significance in the clinically mixed patient groups reviewed here. Of the three whole-brain studies which controlled for clinical state, two found no correlation with diffusion variables (34, 62), and one found a negative correlation between mean FA and scores on the Hamilton Depression Rating Scale (41). Since at least four of the reviewed studies found significant FA differences in remitted euthymic patients, it is likely that abnormalities are not limited to acutely depressed patients. The findings of state-related diffusivity changes should ideally be replicated in longitudinal studies of bipolar patients.

Medication effects, in particular lithium use, may influence gross anatomical findings. In a recent mega-analysis, lithium was found to predict increased volume in hippocampus and amygdala (63). While all papers in our review used patients taking mixed or unspecified medication, Benedetti et al. (34) did a subgroup analysis comparing depressed patients taking lithium only to those taking no medication. They found differences in FA were limited to the lithium group, with no differences in FA between the medication-free group and controls. (Note however that they did find some differences in radial and mean diffusivity in the medication-free group.) No other studies analysed medication-free patients, so the finding is not yet replicated. However, in the three whole-brain studies which controlled for medication use, mood stabilisers were associated with decreased FA in only two clusters (37), there was a non-significant trend towards positive association between lithium dose and FA (41), and there was no

association with lithium dose in the third study (32). The findings suggest a possible, though non-robust, influence of lithium use on diffusivity findings. It is also possible that medication choice and load may correlate with illness severity.

Structural MRI studies in bipolar disorder have suggested an association between illness chronicity and brain changes such as accelerated age-related decline of grey matter volume, increased amygdala volume, and decreased caudate and cerebellar vermis size (64). It is possible that illness chronicity may affect DTI measures as well, although to our knowledge no studies have directly addressed this. The adult studies reviewed here included only patients with chronic illness, with the exception of Chan et al.'s (42) TBSS study and Chen et al.'s (65) VBA study which used first episode remitted and manic patients respectively. Both studies found decreased FA in the posterior projection fibres, in contrast to increased FA in the same fibres found in chronic patients by Versace et al. (37) using TBSS. Chan et al.'s patient group were somewhat atypical however, in that their mean age was similar to patients in other studies despite their first-episode status. Apart from the use of a later-onset bipolar group and differences in substance use history, there are no obvious scanning or methodological differences which may account for the difference in direction of FA findings. This raises the possibility that microstructural WM characteristics, as measured by FA, may change during the course of the illness.

Substance abuse history is also likely to be a significant confounder, as white matter changes on DTI measures have been found in alcoholism (59, 66), abstinent methamphetamine users (67, 68), current methadone users (62), and adolescent cannabis users (69). Though most of our reviewed studies excluded patients with substance abuse histories, four studies included patients with substance abuse histories, which may account for some of the discrepant findings.

Thus, there are many confounding patient factors which may affect DTI results independent of bipolar diagnosis. Subject heterogeneity between and within the reviewed papers may account for poorly replicated results and lack of statistical power, respectively. Future studies would benefit from narrower bipolar phenotypes and larger sample sizes, as has been suggested in reviews of structural MRI in bipolar disorder (64, 70).

LIMITATIONS

We have shown that the available DTI studies in bipolar disorder demonstrate varied and poorly replicated results, with changes in fractional anisotropy involving all major classes of tracts with no predilection for anterior or posterior white matter. Our method has some limitations however.

This systematic review was qualitative rather than quantitative. The important confounder of tract size and architecture is likely to influence labelling by virtue of ease of achieving threshold cluster size, and relative prominence and consistency in atlases leading to labelling bias. These factors are very difficult to correct for statistically, undermining the validity of significance tests. To our knowledge there are no atlases available which report the volume of every candidate white matter tract.

The accurate identification of tracts implicated by clusters of significant voxels in VBA studies is not guaranteed, and has been previously criticised (13, 14). First, clusters do not usually line up neatly with tract edges as portrayed in atlases, but more often cross multiple tracts, making it unclear which tract or tracts are contributing to the abnormality. Second, tracts have intrinsic inter-individual variability, especially as they approach peripheral white matter (45), which is not accounted for by current registration and statistical techniques. There are thus large areas where tract labelling must necessarily be probabilistic. Third, along with inter-individual variability, tract architecture changes as tracts approach the periphery, becoming less coherent and more complex as the fibres spread out towards their ultimate destinations (46, 71). Fourth, a single anatomical tract may contain bundles of separate functional tracts which travel together for part of their course. For example, many association fibers project through the external capsule, and projection fibers pass through the anterior limb of the internal capsule. Similarly, in more anterior WM regions association fibers merge with the projection, thalamic and callosal fibers, and can only be disentangled using recent tractographic techniques (44). Thus, clusters of altered FA in these anatomical tracts or regions cannot be attributed to any specific fibre bundle using current cluster-based methods. Acknowledging the caveats regarding tract identification and implication, we used tract labels as supplied by the authors in their papers, which is likely to be the best estimate, using all information available to them, of which tracts are involved.

Even if we can confidently suggest that a cluster involves part of a tract, it is unclear whether it is reasonable to extrapolate such involvement to the whole tract. Obviously, this depends on how much of

the tract is incorporated in the cluster, which varies greatly between studies. For example, in our review, some clusters implicating anterior projection fibres extended widely from caudal to rostral (62) while some remained circumscribed in central white matter (72). Other studies demonstrate that extrapolating diffusivity measures from cluster to tract gives inconsistent results (13, 72).

TBSS analyses are theoretically better at localising changes to those tracts which form the common white-matter skeleton of most subjects in a group. However, the limitation that an individual anatomic tract may contain multiple functional fibre populations remains. Interestingly, in our review the only clusters which were not labelled came from TBSS studies, suggesting that construction of tract skeletons by TBSS does not always follow labelled tracts from white-matter atlases.

This review was also limited as we did not consider other measures of diffusivity such as mean, radial or axial diffusivity. These measures may illuminate the nature of underlying WM changes contributing to alterations in FA (11). However, the next most commonly reported measure, mean diffusivity or apparent diffusion coefficient, was reported in only five of the fourteen studies reviewed. In addition, the analysis method for non-FA measures was inconsistent, with studies using both voxelwise (38) as well as *a posteriori* ROI analysis of significant clusters (72). Where FA and non-FA measures were both subjected to voxelwise analysis, clusters of altered FA did not necessarily match up with clusters of altered non-FA diffusivity. Bruno et al. (38) suggest that this dissociation between FA and non-FA measures may be due to low resolution with consequent loss of sensitivity for small structures inherent in voxel-based studies. Regardless of the reason, it is unclear how to interpret changes in non-FA diffusivity in the absence of changes in FA. Given that the non-FA results did not appear to be any more consistent than our reviewed FA results, we did not consider that we could generalise from such meagre non-FA data.

In our anterior-posterior analysis, we used the y-coordinates of the peak affected voxels as a proxy measure for the mean location of a cluster. We note that some clusters are large and irregular in shape, and peak voxel is not truly representative of the extent of the cluster. If there were any systematic deviation in peak voxel location relative to the cluster, for example if peak voxels tend to occur in the posterior region of clusters, then our method would be biased. We cannot think of a mechanism which would cause such systematic deviation however.

CONCLUSION

Available whole-brain DTI studies of bipolar disorder in adults produce varied and poorly replicated results. When analysed according to tracts implicated, all major classes of tracts are affected similarly. Using the available published whole-brain DTI data, there is no clear evidence that frontal or limbic WM tracts are affected preferentially in bipolar disorder. Rather, tracts appear to be affected throughout the brain. Additionally, there appears to be no anterior or posterior preference when clusters are considered together. Larger tracts appear to be implicated more frequently, consistent with a generalised diffuse WM abnormality. The anterior thalamic radiation is most often implicated, but only in five out of nine VBA and one out of four TBSS studies.

Different DTI analysis methods, such as VBA, TBSS or tractography, provide different types of anatomical information, and produce different DTI results. Although this precludes a meta-analysis, each analysis method provides useful complementary information which, if used together, may allow additional insight into the location and nature of anatomical changes. Similarly, other DTI measures of diffusivity, such as mD, radD and axiD, may provide complimentary information and facilitate interpretation of FA changes. Future DTI studies in bipolar should attempt to integrate these various measurements and analysis methods into a coherent picture. When considering tract involvement, we suggest that future whole-brain studies attempt to quantify how much of a given tract is affected, and whether proximal or distal ("stem or spray") tract is most affected.

There are many confounders which may contribute independently to alterations in white matter microstructure. The current inconsistencies between studies may be partly attributable to differences between patient groups. Individual studies should strive to use homogenous well-characterised patient groups. Some confounders, such as the effect of illness chronicity, number of affective episodes, and the effect of clinical state, require longitudinal prospective studies to disentangle.

REFERENCES

1. Mann-Wrobel MC, Carreno JT, Dickinson D. Meta-analysis of neuropsychological functioning in euthymic bipolar disorder: an update and investigation of moderator variables. Bipolar Disord. 2011 Jun;13(4):334-42.
2. Emsell L, McDonald C. The structural neuroimaging of bipolar disorder. Int Rev Psychiat. 2009;21(4):297-313.
3. Kempton, al. e. Meta-analysis, Database, and Meta-regression of 98 Structural Imaging Studies in Bipolar Disorder. Arch Gen Psych. 2008;65(9):1017-32.
4. Cerullo MA, al. e. The functional neuroanatomy of bipolar disorder. Int Rev Psychiatry. 2009;21(4):314-22.
5. Phillips ML, Ladouceur CD, Drevets WC. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. Mol Psychiatry. 2008 Sep;13(9):829, 33-57.
6. Mahon K, Burdick KE, Szeszko PR. A role for white matter abnormalities in the pathophysiology of bipolar disorder. Neurosci Biobehav Rev. 2010 Mar;34(4):533-54.
7. Lee CEC, al. e. Normal regional fractional anisotropy and apparent diffusion coefficient of the brain measured on a 3 T MR scanner. Neuroradiology. 2009;51:3-9.
8. Le Bihan D, al. e. Looking into the functional architecture of the brain with diffusion MRI. Nat Rev Neurosci. 2003;4:469-80.
9. Song SK, al. e. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. Neuroimage. 2002;17:1429-36.
10. Sun SW, al. e. Noninvasive detection of cuprizone induced axonal damage and demyelination in the mouse corpus callosum. Magn Reson Med. 2006;55(2):302-8.
11. Beaulieu C. The basis of anisotropic water diffusion in the nervous system - a technical review. NMR Biomed. 2002;15:435-55.
12. Mori S, Jiangyang Z. Principles of Diffusion Tensor Imaging and Its Applications to Basic Neuroscience Research. Neuron. 2006;51:527-39.
13. Kanaan RA, al. e. Tract-specific anisotropy measurements in diffusion tensor imaging. Psychiatr Res: Neuroimaging. 2006;146:73-82.
14. Smith SM, al. e. Tract-based spatial statistics: Voxelwise analysis of multisubject diffusion data. NeuroImage. 2006;31:1487-505.
15. Vederine FE, Wessa M, Leboyer M, Houenou J. A meta-analysis of whole-brain diffusion tensor imaging studies in bipolar disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2011 May 23.
16. Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. Neuroimage. 2002 Jul;16(3 Pt 1):765-80.
17. Sexton K, al. e. A Systematic Review of Diffusion Tensor Imaging Studies in Affective Disorders. Biol Psychiatry. 2009;66:814-23.
18. Heng S, Song AW, Sim K. White matter abnormalities in bipolar disorder: insights from diffusion tensor imaging studies. J Neural Transm. 2010 May;117(5):639-54.
19. Brambilla P, Bellani M, Yeh PH, Soares JC, Tansella M. White matter connectivity in bipolar disorder. Int Rev Psychiatry. 2009;21(4):380-6.
20. Benedetti F, Absinta M, Rocca MA, Radaelli D, Poletti S, Bernasconi A, et al. Tract-specific white matter structural disruption in patients with bipolar disorder. Bipolar Disord. 2011 Jun;13(4):414-24.
21. Lin F, Weng S, Xie B, Wu G, Lei H. Abnormal frontal cortex white matter connections in bipolar disorder: a DTI tractography study. J Affect Disord. 2011 Jun;131(1-3):299-306.
22. Adler CM, Holland SK, Schmithorst V, Wilke M, Weiss KL, Pan H, et al. Abnormal frontal white matter tracts in bipolar disorder: a diffusion tensor imaging study. Bipolar Disord. 2004 Jun;6(3):197-203.
23. Adler CM, Adams J, DelBello MP, Holland SK, Schmithorst V, Levine A, et al. Evidence of white matter pathology in bipolar disorder adolescents experiencing their first episode of mania: a diffusion tensor imaging study. Am J Psychiatry. 2006 Feb;163(2):322-4.
24. Beyer JL, Taylor WD, MacFall JR, Kuchibhatla M, Payne ME, Provenzale JM, et al. Cortical white matter microstructural abnormalities in bipolar disorder. Neuropsychopharmacology. 2005 Dec;30(12):2225-9.
25. Haznedar MM, Roversi F, Pallanti S, Baldini-Rossi N, Schnur DB, Licalzi EM, et al. Fronto-thalamo-striatal gray and white matter volumes and anisotropy of their connections in bipolar spectrum illnesses. Biol Psychiatry. 2005 Apr 1;57(7):733-42.
26. Yurgelun-Todd DA, Silveri MM, Gruber SA, Rohan ML, Pimentel PJ. White matter abnormalities observed in bipolar disorder: a diffusion tensor imaging study. Bipolar Disord. 2007 Aug;9(5):504-12.
27. Wang F, Kalmar JH, Edmiston E, Chepenik LG, Bhagwagar Z, Spencer L, et al. Abnormal corpus callosum integrity in bipolar disorder: a diffusion tensor imaging study. Biol Psychiatry. 2008 Oct 15;64(8):730-3.
28. Macritchie KA, Lloyd AJ, Bastin ME, Vasudev K, Gallagher P, Eyre R, et al. White matter microstructural abnormalities in euthymic bipolar disorder. Br J Psychiatry. 2010 Jan;196(1):52-8.

29. Filley CM. White Matter: Organisation and Functional Relevance. Neuropsychol Rev. 2010;20:158-73.
30. Schmahmann JD, Pandya DN. Fiber Pathways of the Brain. New York: Oxford University Press. 2006.
31. Ha TH, Her JY, Kim JH, Chang JS, Cho HS, Ha K. Similarities and differences of white matter connectivity and water diffusivity in bipolar I and II disorder. Neurosci Lett. 2011 Nov 14;505(2):150-4.
32. Wessa M, Houenou J, Leboyer M, Chanraud S, Poupon C, Martinot JL, et al. Microstructural white matter changes in euthymic bipolar patients: a whole-brain diffusion tensor imaging study. Bipolar Disord. 2009 Aug;11(5):504-14.
33. Lu LH, Zhou XJ, Keedy SK, Reilly JL, Sweeney JA. White matter microstructure in untreated first episode bipolar disorder with psychosis: comparison with schizophrenia. Bipolar Disord. 2011 Nov-Dec;13(7-8):604-13.
34. Benedetti F, Yeh PH, Bellani M, Radaelli D, Nicoletti MA, Poletti S, et al. Disruption of white matter integrity in bipolar depression as a possible structural marker of illness. Biol Psychiatry. 2011 Feb 15;69(4):309-17.
35. Liu JX, Chen YS, Hsieh JC, Su TP, Yeh TC, Chen LF. Differences in white matter abnormalities between bipolar I and II disorders. J Affect Disord. 2010 Dec;127(1-3):309-15.
36. Zanetti MV, Jackowski MP, Versace A, Almeida JR, Hassel S, Duran FL, et al. State-dependent microstructural white matter changes in bipolar I depression. Eur Arch Psychiatry Clin Neurosci. 2009 Sep;259(6):316-28.
37. Versace A, Almeida JR, Hassel S, Walsh ND, Novelli M, Klein CR, et al. Elevated left and reduced right orbitomedial prefrontal fractional anisotropy in adults with bipolar disorder revealed by tract-based spatial statistics. Arch Gen Psychiatry. 2008 Sep;65(9):1041-52.
38. Bruno S, Cercignani M, Ron MA. White matter abnormalities in bipolar disorder: a voxel-based diffusion tensor imaging study. Bipolar Disord. 2008 Jun;10(4):460-8.
39. Oouchi H, Yamada K, Sakai K, Kizu O, Kubota T, Ito H, et al. Diffusion anisotropy measurement of brain white matter is affected by voxel size: underestimation occurs in areas with crossing fibers. AJNR Am J Neuroradiol. 2007 Jun-Jul;28(6):1102-6.
40. Mahon K, Burdick KE, Wu J, Ardekani BA, Szeszko PR. Relationship between suicidality and impulsivity in bipolar I disorder: a diffusion tensor imaging study. Bipolar Disord. 2012 Feb;14(1):80-9.
41. Sussmann JE, Lymer GK, McKirdy J, Moorhead TW, Munoz Maniega S, Job D, et al. White matter abnormalities in bipolar disorder and schizophrenia detected using diffusion tensor magnetic resonance imaging. Bipolar Disord. 2009 Feb;11(1):11-8.
42. Chan WY, Yang GL, Chia MY, Woon PS, Lee J, Keefe R, et al. Cortical and subcortical white matter abnormalities in adults with remitted first-episode mania revealed by Tract-Based Spatial Statistics. Bipolar Disord. 2010 Jun;12(4):383-9.
43. Smith SM, Kindlmann G. Cross-subject comparison of local diffusion MRI parameters. In: Johansen-Berg H, Behrens TEJ, editors. Diffusion MRI: Elsevier; 2009. p. 147-72.
44. Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage. 2007 Jul 1;36(3):630-44.
45. Thiebaut de Schotten M, Ffytche DH, Bizzi A, Dell'Acqua F, Allin M, Walshe M, et al. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. Neuroimage. 2011 Jan 1;54(1):49-59.
46. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex. 2008 Sep;44(8):1105-32.
47. Pfefferbaum A, Adalsteinsson E, Sullivan EV. Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging. Neuroimage. 2005 Jul 1;26(3):891-9.
48. Sullivan EV, Adalsteinsson E, Pfefferbaum A. Selective age-related degradation of anterior callosal fiber bundles quantified in vivo with fiber tracking. Cereb Cortex. 2006 Jul;16(7):1030-9.
49. Wang F, Jackowski M, Kalmar JH, Chepenik LG, Tie K, Qiu M, et al. Abnormal anterior cingulum integrity in bipolar disorder determined through diffusion tensor imaging. Br J Psychiatry. 2008 Aug;193(2):126-9.
50. Snook L, Plewes C, Beaulieu C. Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. Neuroimage. 2007 Jan 1;34(1):243-52.
51. Benazzi F. Mood patterns and classification in bipolar disorder. Curr Opin Psychiatry. 2006 Jan;19(1):1-8.
52. Goes FS, Sanders LL, Potash JB. The genetics of psychotic bipolar disorder. Curr Psychiatry Rep. 2008 Apr;10(2):178-89.
53. Craddock N, O'Donovan MC, Owen MJ. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. J Med Genet. 2005 Mar;42(3):193-204.
54. Bora E, Yucel M, Pantelis C. Neurocognitive markers of psychosis in bipolar disorder: a meta-analytic study. J Affect Disord. 2010 Dec;127(1-3):1-9.
55. Savitz J, van der Merwe L, Stein DJ, Solms M, Ramesar R. Neuropsychological status of bipolar I disorder: impact of psychosis. Br J Psychiatry. 2009 Mar;194(3):243-51.
56. Chaddock CA, Barker GJ, Marshall N, Schulze K, Hall MH, Fern A, et al. White matter microstructural impairments and genetic liability to familial bipolar I disorder. Br J Psychiatry. 2009 Jun;194(6):527-34.
57. McIntosh AM, Job DE, Moorhead TW, Harrison LK, Lawrie SM, Johnstone EC. White matter density in patients with schizophrenia, bipolar disorder and their unaffected relatives. Biol Psychiatry. 2005 Aug 1;58(3):254-7.

58. Hafeman DM, Chang KD, Garrett AS, Sanders EM, Phillips ML. Effects of medication on neuroimaging findings in bipolar disorder: an updated review. Bipolar Disord. 2012 Jun;14(4):375-410.
59. Schulte T, Muller-Oehring EM, Pfefferbaum A, Sullivan EV. Neurocircuitry of emotion and cognition in alcoholism: contributions from white matter fiber tractography. Dialogues Clin Neurosci. 2010;12(4):554-60.
60. Kupferschmidt DA, Zakzanis KK. Toward a functional neuroanatomical signature of bipolar disorder: quantitative evidence from the neuroimaging literature. Psychiatry Res. 2011 Aug 30;193(2):71-9.
61. Darquie A, Poline JB, Poupon C, Saint-Jalmes H, Le Bihan D. Transient decrease in water diffusion observed in human occipital cortex during visual stimulation. Proc Natl Acad Sci U S A. 2001 Jul 31;98(16):9391-5.
62. Liu H, Li L, Hao Y, Cao D, Xu L, Rohrbaugh R, et al. Disrupted white matter integrity in heroin dependence: a controlled study utilizing diffusion tensor imaging. Am J Drug Alcohol Abuse. 2008;34(5):562-75.
63. Hallahan B, al. e. Structural magnetic resonance imaging in bipolar disorder: an international collaborative mega-analysis of individual adult patient data. Biol Psychiatry. 2011;69(4):326-35.
64. Monkul ES, Malhi GS, Soares JC. Anatomical MRI abnormalities in bipolar disorder: do they exist and do they progress? Aust N Z J Psychiatry. 2005 Apr;39(4):222-6.
65. Chen Z, Cui L, Li M, Jiang L, Deng W, Ma X, et al. Voxel based morphometric and diffusion tensor imaging analysis in male bipolar patients with first-episode mania. Prog Neuropsychopharmacol Biol Psychiatry. 2012 Mar 30;36(2):231-8.
66. Harris GJ, Jaffin SK, Hodge SM, Kennedy D, Caviness VS, Marinkovic K, et al. Frontal white matter and cingulum diffusion tensor imaging deficits in alcoholism. Alcohol Clin Exp Res. 2008 Jun;32(6):1001-13.
67. Alicata D, Chang L, Cloak C, Abe K, Ernst T. Higher diffusion in striatum and lower fractional anisotropy in white matter of methamphetamine users. Psychiatry Res. 2009 Oct 30;174(1):1-8.
68. Chung A, Lyoo IK, Kim SJ, Hwang J, Bae SC, Sung YH, et al. Decreased frontal white-matter integrity in abstinent methamphetamine abusers. Int J Neuropsychopharmacol. 2007 Dec;10(6):765-75.
69. Ashtari M, Cervellione K, Cottone J, Ardekani BA, Sevy S, Kumra S. Diffusion abnormalities in adolescents and young adults with a history of heavy cannabis use. J Psychiatr Res. 2009 Jan;43(3):189-204.
70. Kempton MJ, Geddes JR, Ettinger U, Williams SC, Grasby PM. Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. Arch Gen Psychiatry. 2008 Sep;65(9):1017-32.
71. Makris N, Worth AJ, Sorensen AG, Papadimitriou GM, Wu O, Reese TG, et al. Morphometry of in vivo human white matter association pathways with diffusion-weighted magnetic resonance imaging. Ann Neurol. 1997 Dec;42(6):951-62.
72. Mahon K, Wu J, Malhotra AK, Burdick KE, DeRosse P, Ardekani BA, et al. A voxel-based diffusion tensor imaging study of white matter in bipolar disorder. Neuropsychopharmacology. 2009 May;34(6):1590-600.
73. Le Bihan D, al. e. Diffusion Tensor Imaging: Concepts and Applocations. J Magn Res Imaging. 2001;13:534-46.
74. Pavuluri MN, Yang S, Kamineni K, Passarotti AM, Srinivasan G, Harral EM, et al. Diffusion tensor imaging study of white matter fiber tracts in pediatric bipolar disorder and attention-deficit/hyperactivity disorder. Biol Psychiatry. 2009 Apr 1;65(7):586-93.
75. Houenou J, Wessa M, Douaud G, Leboyer M, Chanraud S, Perrin M, et al. Increased white matter connectivity in euthymic bipolar patients: diffusion tensor tractography between the subgenual cingulate and the amygdalo-hippocampal complex. Mol Psychiatry. 2007 Nov;12(11):1001-10.
76. McIntosh AM, Munoz Maniega S, Lymer GK, McKirdy J, Hall J, Sussmann JE, et al. White matter tractography in bipolar disorder and schizophrenia. Biol Psychiatry. 2008 Dec 15;64(12):1088-92.
77. Wang F, Kalmar JH, He Y, Jackowski M, Chepenik LG, Edmiston EE, et al. Functional and structural connectivity between the perigenual anterior cingulate and amygdala in bipolar disorder. Biol Psychiatry. 2009 Sep 1;66(5):516-21.
78. Mori S, van Zijl CM. Fiber tracking: principles and strategies - a technical review. NMR Biomed. 2002;15:468-80.
79. Catani M, Howard RJ, Pajevic S, Jones DK. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. Neuroimage. 2002 Sep;17(1):77-94.
80. Burgel U, Amunts K, Hoemke L, Mohlberg H, Gilsbach JM, Zilles K. White matter fiber tracts of the human brain: three-dimensional mapping at microscopic resolution, topography and intersubject variability. Neuroimage. 2006 Feb 15;29(4):1092-105.
81. Catani M. Diffusion tensor magnetic resonance imaging tractography in cognitive disorders. Curr Opin Neurol. 2006 Dec;19(6):599-606.
82. Ciccarelli O, Parker GJ, Toosy AT, Wheeler-Kingshott CA, Barker GJ, Boulby PA, et al. From diffusion tractography to quantitative white matter tract measures: a reproducibility study. Neuroimage. 2003 Feb;18(2):348-59.
83. Jones DK. Studying connections in the living human brain with diffusion MRI. Cortex. 2008 Sep;44(8):936-52.
84. Buchsbaum MS, Schoenknecht P, Torosjan Y, Newmark R, Chu KW, Mitelman S, et al. Diffusion tensor imaging of frontal lobe white matter tracts in schizophrenia. Ann Gen Psychiatry. 2006;5:19.

Figure 1

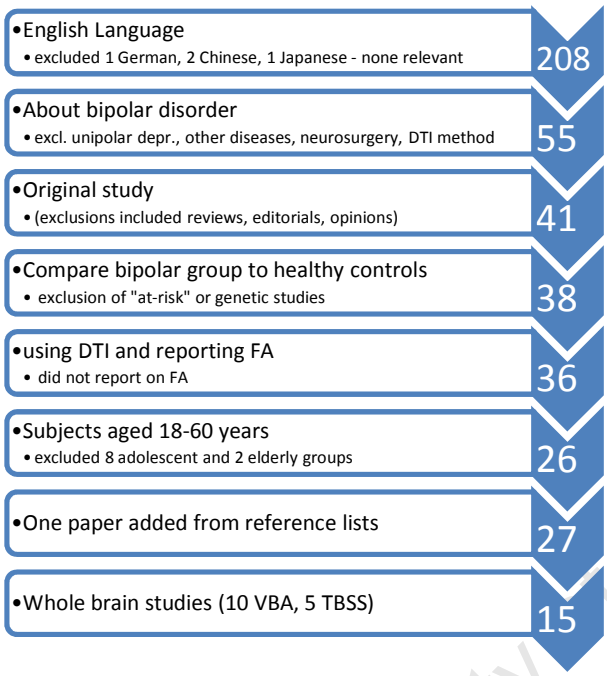


Table 1: Subject characteristics - whole-brain adult DTI studies

Study	Analysis	n	% men	mean age	Diagnosis	Chronicity (years)	Clinically	Meds	% Lith.	Comorbidities / exclusions			
										Psych	Subst	Medical	other
Bruno 2008	VBA	36	36	39	mixed (25 BDI, 11 BDII)	13.8	N/A	mixed	63	nil psych co-morbid	N/A	nil HI, med or neurol hx	
Versace 2008	TBSS	31	35	36	BDI	12	remission & depress combined	mixed	55	nil border-line PD	32% hx subst ab/dep	nil HI, med or neuro hx	
Mahon 2009	VBA/ Tract	30	50	33	mixed (25 BDI, 5 other)	N/A	N/A	N/A	N/A	N/A	30% hx subst ab/dep, unclear if current	nil brain disorder	
Zanetti 2009	VBA	37	35	34	BDI	11.6	depressed / remitted subgroups†	mixed	32	nil border-line PD	37% hx subst ab/dep, nil *current*	nil neuro or HI	
Sussman 2009	VBA	42	52	40	BDI (**75% w psychosis**)	18.3	remission	mixed	57	N/A	N/A	N/A	**+ve fam hx** (1° BD)
Chaddock 2009	VBA	19	47	43	BDI **w hx psychosis**	15.6	remission	mixed	47	1 pt w hx panic only	nil dep in last year, 1 pt w hx alc dep	nil neuro or HI	**+ve fam hx** (1° BD)
Wessa 2009*	TBSS/ VBA	22	50	45	mixed (unspecified)	22	outpatients, **v stable** in remission	mixed	45	1 pt w panic only	nil abuse or dep hx	nil neuro, some vasc risk factors	
Liu 2010	VBA	27	33	35	BDI/BDII subgroups†	8.4	outpatients	mixed	**14**	N/A	nil abuse or dep hx	N/A	
Benedetti 2010	TBSS	40	25	46	BDI	16.4	**depressed**	lithium / no meds subgroups†	65	nil Axis 1	nil abuse or dep hx	nil med or neuro hx	
Chan 2010	TBSS	16	75	37	BDI (**1st episode)**	**0.2**	remission	mixed – **recently started**	37	nil psych hx	nil abuse or dep hx	N/A	
Cui 2011	VBA	18	56	28	BDI	4.8	**manic** in-patients	mixed – 3-7 months non-compliance	N/A	N/A	nil abuse hx	nil neuro, med or HI	**+ve fam hx** (1° or 2° BD)
Ha 2011	VBA	12	25	37	BDI	13.3	remission	mixed	66	N/A	nil abuse hx	nil neuro or signif. med hx	
Chen 2012	VBA	18	100	32	BDI (**1st episode**)	4.2	**manic** in-patients	mixed – 3-7 months non-compliance	N/A	nil Axis 1	nil abuse hx	nil neuro or med hx	**+ve fam hx (**1° BD/MDD)
Mahon 2012	TBSS/ VBA	29	62	34	BDI (14 with suicide att.)	N/A	remission	mixed	N/A	anx. & eating dis. in 12/29	27% subst use disord	N/A	

Atypical subject characteristics compared to other subjects are highlighted in light pink.

N/A: not mentioned in text
HCs: healthy controls
HI: head injury leading to unconsciousness
ab: substance abuse (incl. alcohol)
dep: substance dependence (incl. alcohol)

*Wessa (2009) excluded from tract analysis, included in anterior-posterior analysis.

† Where papers studied subgroups, this review used results from the combined group.

Table 2

Table 2: Scan and analysis characteristics – whole-brain adult DTI studies

Study	Analysis	Reported	n BD vs HC	Tesla	Sequence	Pixel dimensions (mm)	Pixel volume (mm^3)	Directions	b value (sec/mm2)	voxelwise significance (p value)	cluster size threshold (voxels)
Bruno 2008	VBA	FA, mD	36 vs 28	1.5	diffusion- weighted EPI	4 x 4 x 5	80	7	0 - 700	N/A	N/A
Versace 2008	TBSS	FA, radD, axiD	31 vs 25	3.0	single-shot spin EPI	2.5 x 1.6 x 3	12	6	850	$p < 0.001$	5
Mahon 2009	VBA/Tract	FA, radD, axiD	30 vs 38	1.5	spin-echo single-shot EPI	1.7 x 1.7 x 5	13.9	25	1000	$p < 0.001$	50
Zanetti 2009	VBA	FA, mD	37 vs 26	3.0	spin-echo single-shot EPI	1.6 x 1.6 x 3	7.7	6	850	$p < 0.001$	50
Sussman 2009	VBA	FA	42 vs 38	1.5	single-shot spin-echo EPI	N/A	N/A	51	1000	$p < 0.001$	N/A
Chaddock 2009	VBA	FA	19 vs 18	1.5	EPI	2.5 x 2.5 x 2.5	15.6	64	1300	$p < 0.05$	permutation-based method
Wessa 2009	TBSS/VBA	FA, mD	22 vs 21	1.5	EPI	1.9 x 1.9 x 2	7.2	41	700	$P < 0.001$ (TBSS) $p < 0.05$ (VBA)	1 (TBSS) 70 (VBA)
Liu 2010	VBA	FA	27 vs 21	1.5	spin-echo EPI	2.0 x 2.0 x 2.2	8.8	13	900	$p < 0.001$	50
Benedetti 2010	TBSS	FA, mD, radD	40 vs 21	3.0	spin-echo EPI	2.1 x 2.7 x 2.3	13	35	900	TFCE	TFCE
Chan 2010	TBSS	FA, radD, axiD	16 vs 16	3.0	single-shot EPI	2.1 x 2.1 x 3	13.2	15	800	$p < 0.001$	5
Cui 2011	VBA	FA	18 vs 30	3.0	single-shot spin-echo EPI	1.9 x 1.9 x 3	10.5	15	1000	$p < 0.001$	50
Ha 2011	VBA	FA, ADC	12 vs 22	1.5	single-shot spin-echo EPI	2 x 2 x 3	12	15	600	$p < 0.001$	40
Chen 2012	VBA	FA	18 vs 27	3.0	single-shot spin-echo EPI	1.9 x 1.9 x 3	10.5	15	1000	$p < 0.05$[a]	50
Mahon 2012	TBSS	FA	15 vs 15	1.5	spin-echo single-shot EPI	2.5 x 2.5 x 2.5	15.6	21	1000	$p < 0.05$	permutation-based method

TFCE: Threshold-free cluster enhancement
EPI: echoplanar imaging
N/A: not mentioned in text

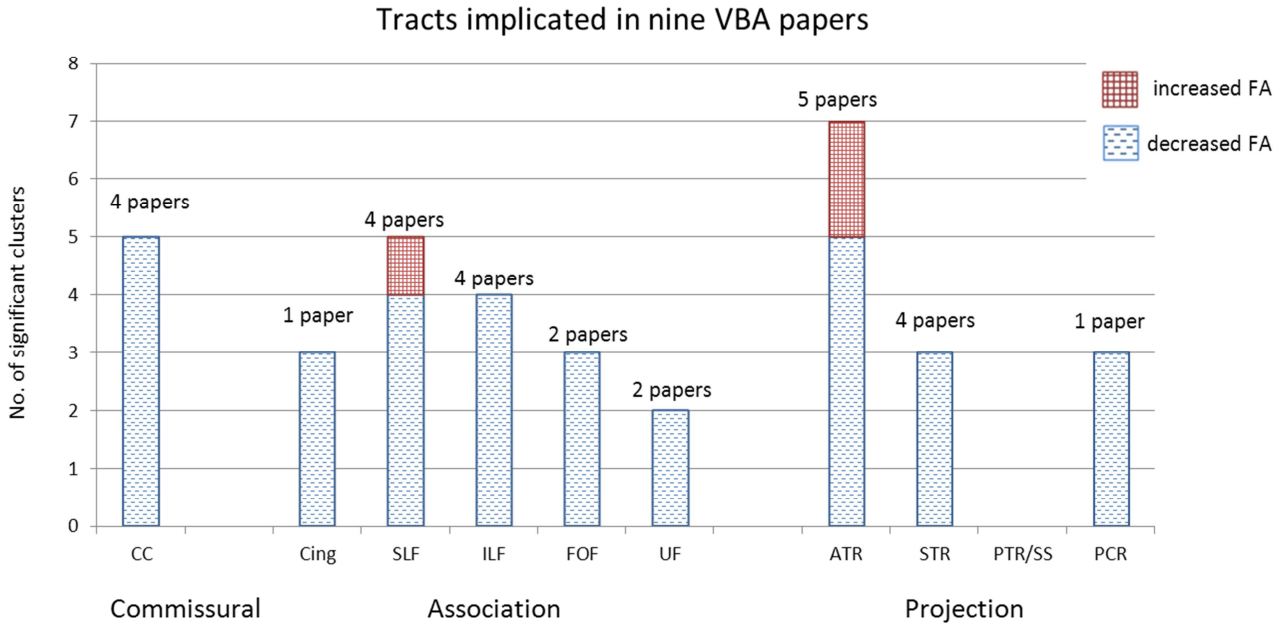
FA: fractional anisotropy
mD: mean diffusivity
axiD, radD: axial diffusivity, radial diffusivity

ADC: apparent diffusion coefficient
a – "false discovery rate" with corrected threshold of $p < 0.05$

Table 4

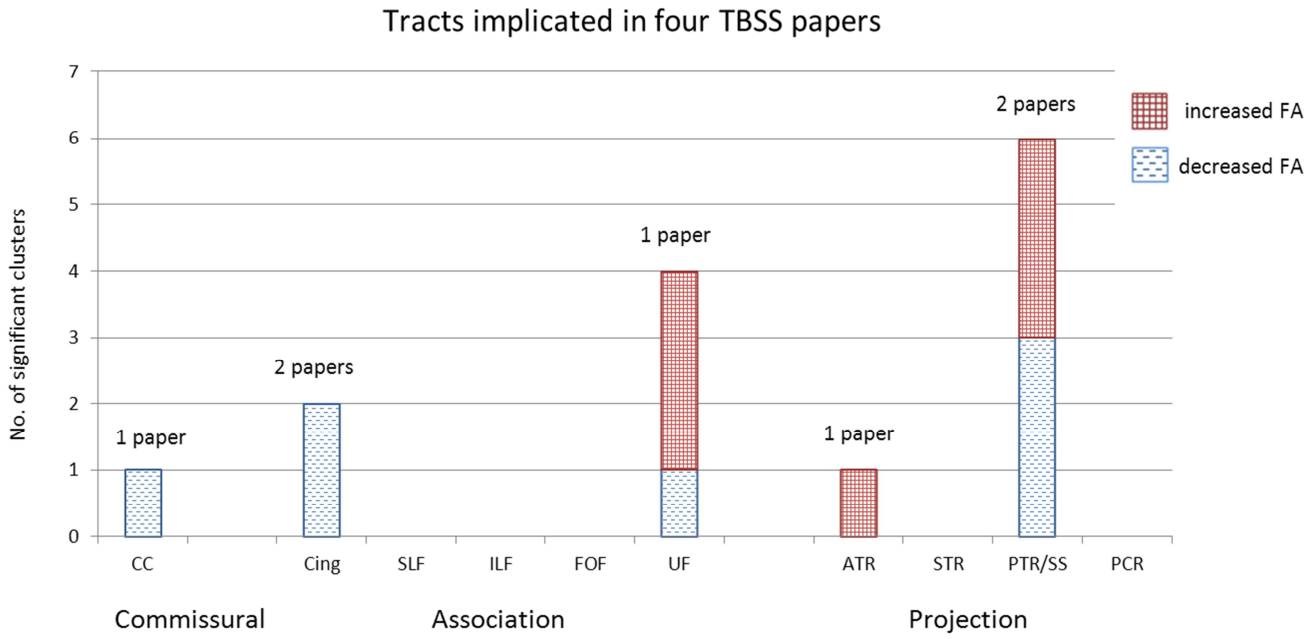
Table 4: Cluster characteristics by analysis method – whole-brain studies						
	No. of studies	Clusters per study (mean,range)	Significant clusters			with tract labels (suitable for frequency analysis)
			decr FA	incr FA	total	
VBA	9	2.83 (1-4)	24*	3*	**27**	27
TBSS	5	4.8, (0-8)	10*	14*	**24**	15
total:	14		34	17	**51**	42
* $p < 0.01$, Fisher's exact						

Figure 2



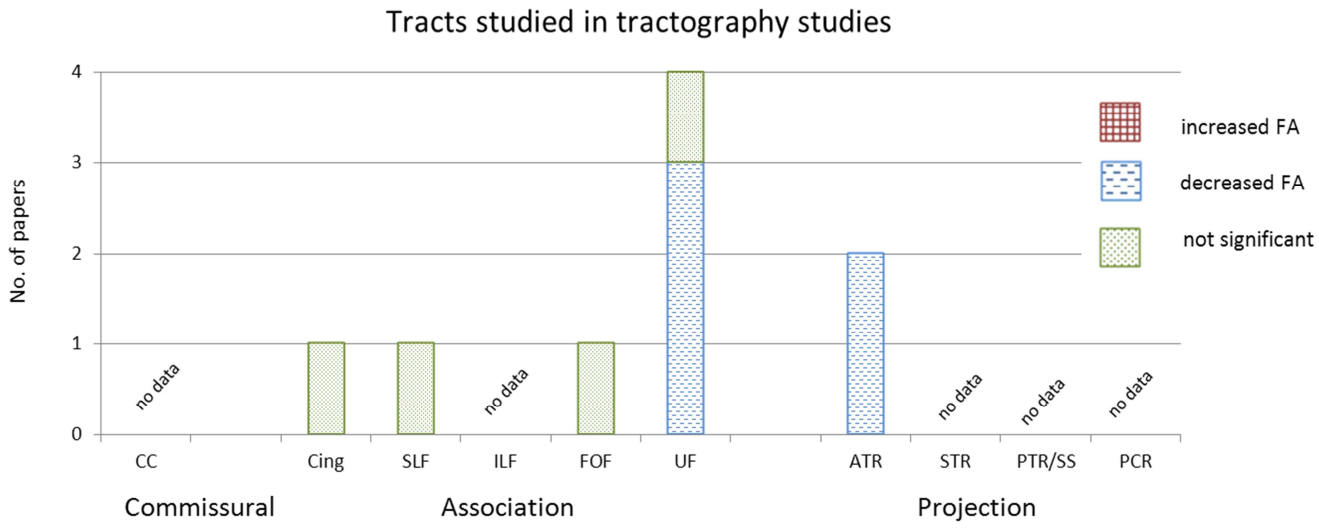
CC: corpus callosum (including forceps minor and major); **Cing**: cingulum; **SLF**: superior longitudinal fasciculus; **ILF**: inferior longitudinal fasciculus; **FOF**: frontooccipital fasciculus; **UF**: uncinate fasciculus; **ATR**: anterior thalamic radiation; **STR**: superior thalamic radiation (including external capsule); **PTR/SS**: posterior thalamic radiation or sagittal stratum; **PCR**: posterior corona radiata. Note: In the whole-brain studies, number of implicated tracts may outnumber number of clusters as clusters could implicate more than one tract.

Figure 3



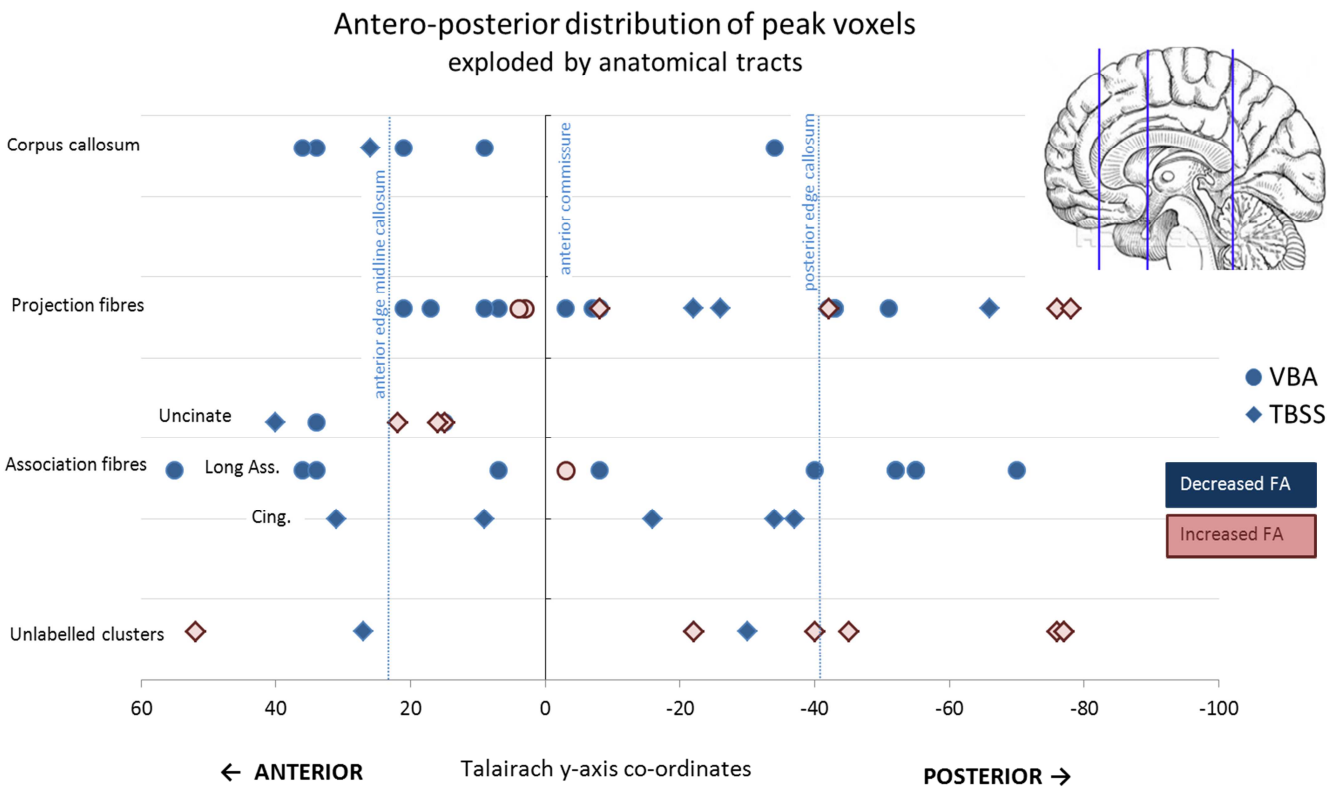
CC: corpus callosum (including forceps minor and major); **Cing**: cingulum; **SLF**: superior longitudinal fasciculus; **ILF**: inferior longitudinal fasciculus; **FOF**: frontooccipital fasciculus; **UF**: uncinate fasciculus; **ATR**: anterior thalamic radiation; **STR**: superior thalamic radiation (including external capsule); **PTR/SS**: posterior thalamic radiation or sagittal stratum; **PCR**: posterior corona radiata. Note: In the whole-brain studies, number of implicated tracts may outnumber number of clusters as clusters could implicate more than one tract.

Figure 4



CC: corpus callosum (including forceps minor and major); **Cing**: cingulum; **SLF**: superior longitudinal fasciculus; **ILF**: inferior longitudinal fasciculus; **FOF**: frontooccipital fasciculus; **UF**: uncinate fasciculus; **ATR**: anterior thalamic radiation; **STR**: superior thalamic radiation (including external capsule); **PTR/SS**: posterior thalamic radiation or sagittal stratum; **PCR**: posterior corona radiata. Note: In the whole-brain studies, number of implicated tracts may outnumber number of clusters as clusters could implicate more than one tract.

Figure 5



Key: Circles and diamonds represent the peak voxels from VBA and TBSS studies respectively. Solid blue indicates decreased FA, shaded pink indicates increased FA.

Note: Where a cluster implicates more than one tract, the peak voxel y-axis coordinate may be plotted more than once under different tracts. For example, the peak voxel at y=36 is plotted both with corpus callosum and with association fibres as both are implicated by that cluster. There are thus six more data points than clusters in the figure.

Supplementary Material

Diffusion Tensor Imaging and Biological Tissues

Technology of DTI

Diffusion Tensor Imaging uses a conventional MRI scanner, though the scan sequence is modified by applying a pair of pulsed magnetic field gradients before each image acquisition. The magnetic field gradients are of the same strength and duration, but with reversed gradients, and separated by sufficient time to allow for diffusion of water. Water molecules are thus effectively "labelled" by the magnetic gradient pulses. The molecules which do not diffuse along the direction of the magnetic gradient are perfectly re-phased by the second pulse and produce a strong MRI signal. However, the molecules which do diffuse along the direction of the magnetic field gradient are imperfectly re-phased by the second pulse, and produce a weaker MRI signal *proportionate to the degree of diffusion*. Thus the signal is sensitive to the degree of diffusion along the direction of the magnetic field gradient. By imaging the same tissue slice repeatedly with magnetic gradients from varying directions, information regarding water diffusion in all three dimensions can be obtained. Diffusion information from at least six gradient directions are required to calculate a *diffusion tensor* for each voxel, representing the rate and direction of diffusion along a primary and two secondary orthogonal axes (73).

The utility of DTI comes from the fact that the normal diffusion of water molecules probes tissue microstructure at a resolution unattainable by any method of direct tissue imaging (73). In tissues where diffusion of water is similar in all directions, such as brain grey matter or cerebrospinal fluid, diffusion is described as relatively *isotropic*. In tissues where diffusion is mostly restricted in one direction, such as white matter or muscle, diffusion is described as *anisotropic*.

DTI measures of diffusivity

The diffusion tensor derived from scanning can be analysed in three main ways to provide qualitatively different but complimentary information on local diffusivity (8). First, *mean diffusivity* characterises the overall average displacement of water molecules in all directions, providing an indication of general barriers to diffusion. It is a scalar value. Second, the degree of anisotropy, usually calculated as *fractional*

anisotropy, provides an indication of the degree to which diffusion may be greater in some directions compared to others. This value gives a clue to the presence and coherence of oriented structures. It is also a scalar value. Third, the *axis of principle diffusion* can be calculated, providing a clue to the direction of orientation of the barriers to diffusion. This value is a vector in three dimensions, and in white matter is usually interpreted as the predominant direction of fibre tracts in a voxel.

Other related values may also be derived from the tensor, and have been used in an attempt to clarify the tissue mechanisms contributing to observed changes in anisotropy in certain conditions, such as bipolar disorder (37, 42, 72). Axial and radial diffusivity reflect the diffusivity parallel to and perpendicular to the direction of the principle axis of diffusion, respectively. These values are thought to reflect diffusion predominantly within axons (axial), verses diffusion predominantly across axons and myelin and in extracellular fluid (radial). There is some evidence in furry rodents that increased radial diffusivity reflects predominantly dysmyelination, and decreased axial diffusivity reflects predominantly axonal damage (9, 10).

The diffusion tensor can be conveniently visualised as a three-dimensional *ellipsoid* shape, representing the degree of projected displacement of water molecules from a point in the tissue. For isotropic diffusion, the ellipsoid is spherical. For more anisotropic diffusion, the ellipsoid may be elongated along one dimension, like a cigar, or flattened along one dimension, like a saucer. Both of these tensor shapes may tell us something about the characteristics of underlying tissue architecture.

In DTI studies, fractional anisotropy (FA) is the most commonly reported measure of diffusivity. It is defined as a ratio of sum-of-squared differences in diffusivity in three orthogonal directions, over the sum-of-squares of diffusivity in the same three orthogonal directions, and represents the proportion of total diffusivity attributable to anisotropy. As such, it varies from 0 (complete isotropy, as may be expected in cerebrospinal fluid) to 1 (complete anisotropy). In practice, FA in the brain varies from about 0.2 (in grey matter) to 0.8 (in parts of white matter) (7).

Tissue contributions to anisotropy

Fractional anisotropy in white matter, as measured by DTI, is dependent both on the microscopic cellular structure of individual axons, as well as on the macroscopic tissue structure of bundles of axons

together. The exact mechanisms and relative contributions of cellular and extracellular structures to diffusion impedance and anisotropy remain unclear (8), but certain empirical observations are relevant. On a microscopic cellular level, the majority of anisotropy is contributed by the axonal cell membrane, with the myelin sheath probably playing an important modulating role (8). That is, unmyelinated axons compared to myelinated axons typically demonstrate only a 20% decrease in measured anisotropy (11). Axon diameter also plays a role when comparing axons of widely differing diameter (11), though the relevance of this to WM in the brain is unclear. Other potential cellular barriers to diffusion, such as microtubules and intracellular organelles contribute minimally to anisotropy. Active axonal transport similarly has little effect on measured longitudinal diffusion. Experimental evidence thus indicates that it is the *spatial organisation of membranes* which primarily contribute to anisotropy (8).

In addition to diffusion anisotropy on the microscopic level, there must be a degree of coherence and directionality in the macroscopic structure of the tissue for anisotropy to be measured in a voxel. An individual voxel in DTI typically has volume of 8-27 mm^3, thus incorporating thousands of axons and extracellular material. Measured diffusion is an average over all these elements. It is only the regular coherent arrangement of axons into bundles of parallel fibres which permits the measurement of anisotropy by DTI (12). Changes in measured FA may thus be due to changes in fibre coherence, for example where WM tracts cross each other. Increased extracellular space, or lower fibre density, may also contribute to decreased anisotropy (34).

Differences in scanning parameters interact with tissue architecture to influence measured FA. Larger volume voxels, for example, will tend to contain more heterogeneous fibre populations than smaller voxels. Since measured FA is a measure of average diffusion in a voxel, larger voxels will tend to report lower FA than smaller voxels for equivalent tissue. Smaller voxels will tend to contain more coherent fibre populations, so a higher resolution scan may produce higher measured FA values for exactly equivalent tissues (12). This difference will be most pronounced in areas of architectural complexity. Thus FA values are dependent on voxel size, which in turn is related to scanning parameters and magnet strength. As a corollary, the type of tissue changes likely to be picked up by DTI studies between groups is sensitive to voxel volume. Scans with larger voxels may be more dependent on fibre coherence for measured FA, while scans with smaller voxels may be more dependent on axonal integrity and myelination for FA.

In summary, DTI does not distinguish between microscopic cellular and macroscopic structural contributions to measured FA, so a clear biological interpretation of changes in FA can be difficult. Nevertheless, changes in FA are interpreted as reflecting some alteration in WM anatomy, whether it be due to neuronal injury or changes in myelination, axon diameter or fibre coherence. In some instances where neuronal architecture is likely to be constant, relative changes in radial and axial diffusivity may give an indication of the nature of the underlying neuronal pathology (9, 10). The routine interpretation of FA as a measure of WM integrity is thus an oversimplification (12).

DTI Analysis Methods

DTI brain images have been analysed in various ways, each with their own strengths and weaknesses.

ROI Methodology and Limitations

Early DTI studies of bipolar patients from 2004 used a "region-of-interest" (ROI) approach, in which the *mean* FA value (or other measure of diffusivity) is measured over a circumscribed area of white matter chosen in advance. There are many ways of selecting and placing the ROI, some of which may isolate a predefined geometric region (circle or cube, for example), and some of which may attempt to isolate parts of specific tracts. ROI's may be delineated manually by tracing or measuring, or automated according to anatomical landmarks. The more sophisticated ROI technique using tractography is described later.

After the mean FA in the ROI is calculated for each subject, values are averaged to provide a group-mean FA value which can be compared to another group using Student T-test.

ROI approaches are statistically simple, and regions can be chosen a priori based on their hypothesised relation to, and involvement in, the disorder of interest. The rest of the brain volume, however, remains unexamined, and it is impossible to know if any significant findings are limited to that region or more widespread.

Although conceptually simple, more recent work has shown that results from manual ROI placement are somewhat unreliable due to dependence on subtle methodological differences (13, 15). Apparently minor changes in shape, size, or placement of the ROI may lead to different and even conflicting results, even in a relatively easily identified and apparently homogenous tract like the genu of the corpus callosum (13).

In complex areas of white matter, such as the frontal lobe or corona radiata, it may be impossible to identify which tracts impinge on a ROI, and whether they are coherent or crossing. In general, normal variability in tract anatomy will make it difficult to consistently isolate individual tracts across subjects, especially when ROI's are placed by measuring from landmarks rather than underlying WM anatomy. These problems are exacerbated by the smoothing of diffusivity images, which lowers resolution and increases partial volume effects. These problems of contamination, tract identification and partial volume effects may be minimised by placing the ROI in the centre, or stem, of well-defined tracts (14).

Two approaches are commonly used to define the ROI in bipolar DTI studies, both attempting to isolate equivalent WM regions in an reliable objective manner. The first approach uses measurements from constant anatomical landmarks, such as the anterior commissure, to define a geometrically shaped ROI in white matter (22, 23). In some cases stereotaxic co-ordinates are used to place the ROI (25). The second approach uses the subject's reference diffusion- or T1-weighted MRI image to place the geometrically-shaped ROI directly in a specified white matter tract (24, 26, 28, 74). The third approach uses manual tracing of colour-coded tensor maps to define a ROI which follows the contours of a tract like the corpus callosum or cingulum (27, 49). Many studies report good reliability using their method of ROI definition, but methodological differences should prompt caution when comparing results between studies. These differences in methodology may partly explain the mixed and sometimes conflicting results.

Tractography methodology and limitations

An alternate method of overcoming the traditional ROI difficulties of consistently and accurately localising tracts, is to identify the whole tract using automated DTI *fibre-tracking* (21, 75-77). The hope is that the statistical simplicity of ROI analysis can be retained while adding improved tract isolation and

consideration of tracts along more of their length. Since tractography does not require inter-subject image registration, the problems of misregistration are avoided.

Fibre-tracking makes use of the *principle diffusion axis* derived from the diffusion tensor, to estimate the predominant direction of the fibres in each voxel. Starting from a seed ROI or voxel, the tract is incrementally plotted from voxel to voxel by following the principle diffusion axis (78, 79), until the fibre direction changes abruptly (more than a specified angle) or the FA of the voxel becomes less than a specified value, typically 0.2. These termination criteria ensure that only voxels which are likely to be part of the tract are included in the ROI mask. Usually the tract is defined as including all fibres which pass through two seed-regions of interest, such as orbital prefrontal cortex and anterior temporal cortex if one wanted to track the uncinate fasciculus (Houenou 2007, McIntosh 2008, Wang 2009, Lin 2010). White-matter anatomy as delineated by DTI tractography has been validated against gross anatomy as demonstrated by traditional dissection (44). However, it should be noted that there is demonstrable inter-individual variability in tract architecture, more evident in extended areas of tracts, such as the thalamic radiations (45, 80).

Tractography defines the ROI more reliably than traditional ROI methods (81), though still depends critically on the placement of seed-regions in relation to surrounding architectural complexity (13, 82). Tractographic ROI's are larger than and have less variance of intra-group mean FA's than smaller traditional ROI's, allowing for more statistical power. Notably, because the tract is considered along most of its length (until FA falls below 0.2), and includes regions where fibres begin to fan out and lose coherence, the mean FA from a tractographic study is likely to be less than that from a traditional study where the ROI may be placed in the more coherent "stem" of a tract (13). The interaction between tract architecture, measured FA, and fibre-tracking termination conditions suggests a potential bias towards including the regions of tracts which are less complex and more coherent, while excluding regions of tracts which are less coherent with lower measured FA. Fibre-tracking may end prematurely due to decreased FA if fibres become more dispersed or are contaminated by crossing fibres. Fibre-tracking may be inaccurate especially in areas where populations of differently oriented fibres cross, leading to falsely reconstructed fibres, as well as false negatives when thinly spread fibres diverge from the main tract (82, 83). If there is a systematic difference between groups in tract architecture (as may be due to inter-group differences in tract coherence or number of crossing fibres), this may bias results as less of the dispersed tracts will be included in the ROI. It is thus necessary in tractography studies to check that

there is no systematic difference in the volume and length of tract masks between groups (13, 44, 84). Similarly, tractography is more susceptible than traditional ROI to partial-volume effects and biases dependent on voxel size (as described above) since it includes areas of more thinly spread fibres.

Voxel-based analyses

Traditional voxel-based analyses: methodology and limitations

Voxel-based analyses (VBA) consider the whole brain, performing a voxel-wise comparison between subject images and control images. This operator-independent process is unbiased regarding a priori assumptions about affected areas, and is particularly suitable when changes may be diffuse or where there is no clear a priori hypothesis (13). Statistical analysis is more complicated than with ROI studies however, as the large number of voxel-wise comparisons must be corrected for. This is usually done by setting a high threshold of statistical significance for each voxel. In addition a specified number of contiguous voxels is required to reach statistical significance, so controlling for isolated spurious voxels due to noise. These "clusters" are typically thresholded to consist of at least 50-100 contiguous significant voxels.

Voxel-based analyses rely on exact registration and normalization of the diffusivity images, as corresponding voxels in all images from both groups should represent exactly the same area of anatomy (14). Tracts of white matter are potentially finer and more complex in structure than the gross anatomy which normalization techniques were originally optimized for. Current techniques for normalisation of relatively low-resolution diffusivity images remain inadequate (21), and misregistration artefacts are likely. This may have two effects, depending on the complexity of the underlying anatomy. In less complex anatomical areas, any systematic differences in registration of white matter tracts between subject groups, as may occur because of volumetric or shape structural differences between groups, will cause an artefactual margin of increased statistical significance along misaligned high contrast borders. In more complex anatomical areas with narrow fibre bundles, misregistration across images will tend to minimise any significant differences between groups. Both effects are exaggerated by the smoothing which is applied to diffusivity images. These misregistration artefacts complicate interpretation of VBA findings. (14).

Finally, accurate localisation of differences to specific tracts is difficult with VBA (13). Clusters identified by VBA methods seldom align neatly with white matter tracts, and may include parts of more than one tract. In areas where tracts cross each other and the architecture is complex, a single cluster will include mixed fibres from multiple tracts. Further difficulties in identifying affected tracts arise from the relatively low resolution of smoothed diffusivity images, coupled with normal anatomical variability. Resolving a cluster into component tracts is further hampered by relatively low detail in white-matter atlases (13).

Tract-based spatial statistics (TBSS): methodology and limitations

Tract-based spatial statistics (TBSS) aims to combine the strengths of tractography and voxel-based whole-brain approaches, while avoiding their weaknesses (14). Thus, it aims to reliably identify major white matter tracts and avoid partial volume and registration artefacts, while considering the whole-brain in an unbiased operator-independent manner.

TBSS makes use of the simplifying assumption that it is usually only the centre of WM tracts which we are interested in from a DTI perspective, and the FA values of tract peripheries can be ignored. Thus, the mask which we are interested in is effectively the central "skeleton" of all the white matter tracts. This skeleton is derived from an average FA image over all the subjects, after registration. Exact registration is not needed, as areas of poor registration between subjects will have high variance in the averaged image, and thus not form part of the skeleton. Also, smoothing is not necessary, which avoids partial-volume effects. The white-matter skeleton is extracted from the averaged FA image by "non-maximum suppression", leaving only the middle part of the tracts with highest average FA (14). Once this average skeleton template has been derived, the WM tracts in each subject's images are projected perpendicularly onto the nearest part of the average WM skeleton. Thus one ends up with the centre of each tract in each subject's diffusivity image perfectly aligned to an average WM skeleton. Peripheral non-maximal voxels of the WM tract have been discarded. Since all subject images are perfectly aligned, a voxelwise comparison between subject groups is valid, using the same statistical technique as in VBA analyses.

TBSS is useful for detecting alterations in FA in the middle of tracts where FA is highest. However, it is not able to investigate areas of WM peripheral to the larger tracts, where there is more likely to be fibre incoherence or crossing. However, it is possible that even central parts of WM tracts in one group may contain more crossed or incoherent fibres than the central parts from the comparison group. Therefore one cannot automatically assume that FA alterations are exclusively due to differences in myelination or axonal integrity.

Summary of limitations of DTI analysis methods

Each method of analysis has distinct limitations when localising changes in diffusivity to specific tracts. Traditional ROI studies frequently struggle to isolate tracts, are extremely sensitive to issues of ROI placement, and the results are seldom replicated. Tractography studies isolate individual tracts, but cannot tell if altered diffusivity is spread throughout the tract or localised to certain regions. In addition, they cannot place the findings in the context of the rest of the brain as they examine only specific regions at a time. VBA studies consider the whole brain at once and give insight into distribution of diffusivity alterations, but are subject to misregistration artefacts and struggle to accurately identify affected tracts. TBSS studies overcome many of these difficulties, but are limited to examining the "skeleton" of only the largest most anatomically consistent tracts. Hybrid methods, for example identifying significant clusters followed by tractography (72), provide useful insights but have yet to be replicated.